

# EVALUATING DEFAULT PRIORS WITH A GENERALIZATION OF EATON'S MARKOV CHAIN

BRIAN P. SHEA AND GALIN L. JONES  
SCHOOL OF STATISTICS  
UNIVERSITY OF MINNESOTA

**ABSTRACT.** We consider evaluating improper priors in a formal Bayes setting according to the consequences of their use. Let  $\Phi$  be a class of functions on the parameter space and consider estimating elements of  $\Phi$  under quadratic loss. If the formal Bayes estimator of every function in  $\Phi$  is admissible, then the prior is strongly admissible with respect to  $\Phi$ . Eaton's method for establishing strong admissibility is based on studying the stability properties of a particular Markov chain associated with the inferential setting. In previous work, this was handled differently depending upon whether  $\varphi \in \Phi$  was bounded or unbounded. We introduce and study a new Markov chain which allows us to unify and generalize existing approaches while simultaneously broadening the scope of their potential applicability. To illustrate the method, we establish strong admissibility conditions when the model is a  $p$ -dimensional multivariate normal distribution with unknown mean vector  $\theta$  and the prior is of the form  $\nu(\|\theta\|^2)d\theta$ .

## 1. INTRODUCTION

Suppose we are in a parametric setting, and we are considering use of an improper prior measure that yields a proper posterior distribution. Such priors arise in the absence of honest prior belief about parameter values and are typically derived from structural arguments based on the likelihood or the parameter space [17]. Thus, an improper prior, rather than being a statement of beliefs specific to a situation, is a default. Such priors, proposed from likelihood or invariance arguments, require evaluation, just as estimators proposed from likelihood or invariance arguments require evaluation, and

---

*Date:* September 7, 2011.

*Key words and phrases.* Admissibility, Improper prior distribution, Symmetric Markov chain, Recurrence, Dirichlet form, Formal Bayes rule.

Brian Shea (bshea@stat.umn.edu) is a graduate student and Galin Jones (galin@umn.edu) is an Associate Professor. Research supported by the National Science Foundation and the National Institutes of Health. The authors thank Morris Eaton for many helpful conversations.

an attractive avenue is to evaluate the prior according to the consequences of its use. That is, we can evaluate the prior by examining properties of the resulting posterior inferences. The criterion we use to judge posterior inferences is known as *strong admissibility*. This concept was introduced by Eaton [6] and has given rise to a substantial theory [6, 7, 8, 9, 10, 11, 12, 13, 14, 18]. Our goal in the rest of this section is to convey the basic idea behind strong admissibility and the way it is studied here. We also summarize our main results without delving too far into the details, which are dealt with carefully later.

Suppose the sample space  $\mathcal{X}$  is a Polish space with Borel  $\sigma$ -algebra  $\mathcal{B}$  and the parameter space  $\mathfrak{V}$  is a Polish space with Borel  $\sigma$ -algebra  $\mathcal{C}$ . Let  $\{P(\cdot|\theta), \theta \in \mathfrak{V}\}$  be a family of sampling distributions where we assume that for each  $B \in \mathcal{B}$ ,  $P(B|\cdot)$  is  $\mathcal{C}$ -measurable and for each  $\theta \in \mathfrak{V}$ ,  $P(\cdot|\theta)$  is a probability measure on  $(\mathcal{X}, \mathcal{B})$ . Let  $\nu$  be a  $\sigma$ -finite measure on the parameter space  $\mathfrak{V}$  with  $\nu(\mathfrak{V}) = \infty$ . Throughout the marginal on  $\mathcal{X}$

$$(1) \quad M(dx) := \int_{\mathfrak{V}} P(dx|\theta) \nu(d\theta) .$$

is assumed to be  $\sigma$ -finite. In this case, the disintegration

$$(2) \quad Q(d\theta|x) M(dx) = P(dx|\theta) \nu(d\theta)$$

generalizes Bayes theorem and implicitly defines formal posterior distributions  $Q(\cdot|x)$  on the parameter space. Note that for each  $x \in \mathcal{X}$ ,  $Q(\cdot|x)$  is a probability measure on  $(\mathfrak{V}, \mathcal{C})$  and for each  $C \in \mathcal{C}$ ,  $Q(C|\cdot)$  is  $\mathcal{B}$ -measurable. Taraldsen and Lindqvist [23] provide a recent, accessible introduction to the existence of formal posterior distributions while one can consult Eaton [5, 6] and Johnson [16] for more details and references.

Suppose  $\varphi : \mathfrak{V} \rightarrow \mathbb{R}^p$  for  $p \geq 1$ , and consider estimating  $\varphi(\theta)$ . The formal Bayes estimator of  $\varphi$  under squared error loss is the posterior mean

$$(3) \quad \hat{\varphi}(x) := \int_{\mathfrak{V}} \varphi(\theta) Q(d\theta|x) .$$

Let  $\|\cdot\|$  denote the usual Euclidean norm. If  $\delta(x)$  is any estimator of  $\varphi(\theta)$ , the risk function of  $\delta$  is

$$(4) \quad R(\delta; \theta) := \int_{\mathcal{X}} \|\varphi(\theta) - \delta(x)\|^2 P(dx|\theta) .$$

The estimator  $\delta$  is *almost- $\nu$  admissible* if for any other estimator  $\delta_1$  such that  $R(\theta; \delta_1) \leq R(\theta; \delta)$  for all  $\theta \in \mathfrak{V}$ , then the set  $\{\theta : R(\theta; \delta_1) < R(\theta; \delta)\}$  has  $\nu$ -measure 0.

Since we will use admissibility to judge the prior, our interests are more ambitious than establishing admissibility of a single estimator. Let  $\Phi$  be a class of functions defined on the parameter space. If the formal Bayes estimator of every  $\varphi \in \Phi$  is almost- $\nu$  admissible, then we say the prior (equivalently the posterior) is *strongly admissible with respect to  $\Phi$* . A prior is strongly admissible if it is robust against risk dominance within the class  $\Phi$ . Since a default prior will undergo repeated use, it is important for the range of appropriate uses to be clearly defined and desirable that the range be as large as possible. We can then endorse the improper prior insofar as it avoids unreasonable actions in a variety of such problems.

Previous work on strong admissibility focused on the case where  $\Phi$  consisted of a single unbounded function [2, 3, 8], or all bounded functions [6, 7, 9, 10, 11, 12, 13, 18]. In either case, strong admissibility was established either by verifying sufficient conditions for the admissibility of an estimator established by Brown [4] or via Markov chain arguments using an approach developed by Eaton [6, 8]. We study the latter method.

Eaton's method for establishing almost- $\nu$  admissibility of formal Bayes estimators is based on the recurrence properties of a Markov chain associated with the inferential setting; the relevant notion of recurrence is defined in the next section. However, different Markov chains were required depending upon whether  $\varphi$  was bounded [6] or unbounded [8]. We introduce and study a new Markov chain which allows us to unify and generalize these existing approaches while simultaneously broadening the scope of their potential applicability. The expected posterior

$$(5) \quad R(d\theta|\eta) = \int_{\mathcal{X}} Q(d\theta|x) P(dx|\eta) .$$

is a Markov kernel on  $\boldsymbol{\vartheta}$ . We study transformations of  $R$ , which are now described. Let  $f : \boldsymbol{\vartheta} \times \boldsymbol{\vartheta} \rightarrow \mathbb{R}^+$  satisfy  $f(\theta, \eta) = f(\eta, \theta)$  and set

$$\mathcal{T}(\eta) = \int_{\boldsymbol{\vartheta}} f(\theta, \eta) R(d\theta|\eta) .$$

If  $0 < \mathcal{T}(\eta) < \infty$  for all  $\eta \in \boldsymbol{\vartheta}$ , then

$$T(d\theta|\eta) = \frac{f(\theta, \eta)}{\mathcal{T}(\eta)} R(d\theta|\eta)$$

is a Markov kernel on  $\boldsymbol{\vartheta}$ . Recurrence of the Markov chain associated with  $T$  implies the almost- $\nu$  admissibility of formal Bayes estimators with respect to a large class of functions. Define

$$(6) \quad \Phi_f := \{ \varphi : \|\varphi(\theta) - \varphi(\eta)\|^2 \leq M_\varphi f(\theta, \eta) \text{ some } 0 < M_\varphi < \infty \} .$$

We prove that if the Markov chain defined by  $T$  is recurrent, then the formal Bayes estimator of every function in  $\Phi_f$  is almost- $\nu$  admissible, and we say the prior  $\nu$  is *strongly admissible with respect to  $\Phi_f$* . The following example illustrates this technique.

**Example 1.** Let  $X$  be a  $p$ -dimensional normal random variable with identity covariance  $I_p$  and unknown location  $\theta$ . Let  $p$ -dimensional Lebesgue measure  $\nu(d\theta) = d\theta$  be our improper prior. The proper posterior for  $\theta$  is a normal with mean  $x$ —the observed value of the random variable  $X$ —and covariance  $I_p$ . The kernel

$$R(d\theta|\eta) = \int_{\mathcal{X}} Q(d\theta|x)P(dx|\eta) = (4\pi)^{-p/2} \exp\left(-\frac{1}{4}\|\theta - \eta\|^2\right) d\theta$$

describes a  $N(\eta, 2I_p)$  random variable. Let  $d$  be an arbitrary positive constant and note that  $f(\theta, \eta) = \|\theta - \eta\|^2 + d$  is symmetric and uniformly bounded away from 0. Further

$$\mathcal{T}(\eta) = \int f(\theta, \eta) R(d\theta|\eta) = \int (\|\theta - \eta\|^2 + d) R(d\theta|\eta) = 2p + d.$$

Thus,

$$T(d\theta|\eta) = \frac{\|\theta - \eta\|^2 + d}{2p + d} R(d\theta|\eta)$$

and the chain with kernel  $T$  is a random walk on  $\mathbb{R}^p$ . Note that

$$\int \|\theta - \eta\|^p T(d\theta|\eta) < \infty$$

since a normal distribution has moments of all orders and hence for  $p = 1$  or  $p = 2$ , the chain is recurrent [1, 20]. We conclude that for  $p = 1$  or  $p = 2$  Lebesgue measure is strongly admissible. That is, the formal Bayes estimators of all functions  $\varphi$  satisfying

$$\|\varphi(\theta) - \varphi(\eta)\| \leq M_\varphi (\|\theta - \eta\|^2 + d) \quad \text{for some } 0 < M_\varphi < \infty$$

are almost- $\nu$  admissible. Since  $d$  is arbitrary this includes all bounded functions as well as many unbounded functions.

Lebesgue measure is not strongly admissible with respect to  $\Phi_f$  when  $p \geq 3$  since the James-Stein estimator dominates the formal Bayes estimator of  $\varphi(\theta) = \theta$ . In Section 4 we consider the normal means problem with an alternative prior when  $p \geq 3$ . We will return to this example below.

Recurrence of the Markov chain described by  $T$  implies more than we have so far claimed. Suppose the function  $u: \mathfrak{V} \rightarrow \mathbb{R}^+$  is bounded away from zero and infinity so that  $1/c < u(\theta) < c$  for some constant  $c > 1$  and every  $\theta \in \mathfrak{V}$ . The measure defined by

$$(7) \quad \nu_u(d\theta) := u(\theta) \nu(d\theta)$$

is a *bounded perturbation* of  $\nu$ . Let  $\mathcal{F}_\nu$  be the family of all bounded perturbations of  $\nu$ . Observe that  $\nu \in \mathcal{F}_\nu$  and that the other elements of  $\mathcal{F}_\nu$  are measures with tail behavior similar to  $\nu$ . Eaton [6] showed that recurrence of the chain with kernel  $R$  implied the formal Bayes estimators of any bounded function is almost- $\nu$  admissible for every prior in  $\mathcal{F}_\nu$ . We extend this result and show that recurrence under  $T$  is sufficient for the strong admissibility with respect to  $\Phi_f$  of every element of  $\mathcal{F}_\nu$ . In this case, we say the family  $\mathcal{F}_\nu$  is *strongly admissible with respect to  $\Phi_f$* .

**Example 2.** Recall the setting of Example 1. Let  $\mathcal{F}_\nu$  be the bounded perturbations of  $p$ -dimensional Lebesgue measure  $\nu(d\theta) = d\theta$ . That is, elements of  $\mathcal{F}_\nu$  are measures of the form  $\nu_u(d\theta) = u(\theta)d\theta$  where  $1/c < u(\theta) < c$  for some constant  $c > 1$ . Recall that  $\Phi_f$  is the class of all bounded functions and all functions satisfying

$$\|\varphi(\theta) - \varphi(\eta)\| \leq M_\varphi (\|\theta - \eta\|^2 + d) \quad \text{for some } 0 < M_\varphi < \infty.$$

The recurrence of the Markov chain governed by  $T$  when  $p = 1$  or  $p = 2$  implies the formal Bayes estimators of every  $\varphi \in \Phi_f$  are almost- $\nu_u$  admissible for every prior  $\nu_u \in \mathcal{F}_\nu$ . That is,  $\mathcal{F}_\nu$  is strongly admissible with respect to  $\Phi_f$ .

Our main results generalize existing work [6, 8] and in fact unify the analysis for bounded and unbounded functions. We show that the Markov kernel  $R$  can be transformed to define many Markov chains, any one of which might be used to demonstrate strong admissibility, thus greatly broadening the scope of potential applications. We are never concerned narrowly with a single admissibility problem but broadly with species of problems. Moreover, solving a single representative problem, which representative we are free to elect, solves any problem within a bounded rate of change—whether of the function to be estimated or the prior used to estimate it.

The remainder is organized as follows. Section 2 gives some background on recurrence for general state space Markov chains. Section 3 presents the main results and Section 4 illustrates the main results by considering the multivariate normal means problem. Finally, many technical details are deferred to the appendices.

## 2. RECURRENCE OF MARKOV CHAINS

The goal of this section is to introduce a general notion of recurrence for Markov chains. Let  $\mathcal{W}$  be a Polish space and denote the Borel  $\sigma$ -algebra by  $\mathcal{D}$ . Let  $K : \mathcal{D} \times \mathcal{W} \rightarrow [0, 1]$ . Then  $K$  is a *Markov transition kernel* on the measurable space

$(\mathcal{W}, \mathcal{D})$  if  $K(D|\cdot)$  is a nonnegative measurable function for every  $D \in \mathcal{D}$  and  $K(\cdot|w)$  is a probability measure for every  $w \in \mathcal{W}$ .

The kernel  $K$  determines a time-homogeneous Markov chain  $W = \{W_0, W_1, W_2, \dots\}$  on the product space  $\mathcal{W}^\infty$  which is equipped with the product  $\sigma$ -algebra  $\mathcal{D}^\infty$ . Note that conditional on  $W_n = w$ , the law of  $W_{n+1}$  is  $K(\cdot|w)$ . Given  $W_0 = w$ , let  $\Pr(\cdot|w)$  be the law of  $W$  on  $\mathcal{W}^\infty$ .

Suppose  $D \in \mathcal{D}$ . The random variable

$$\tau_D = \begin{cases} \infty & \text{if } W_n \notin D \text{ for all } n \geq 1, \text{ or} \\ \text{the smallest } n \geq 1 \text{ such that } W_n \in D & \text{otherwise} \end{cases}$$

is a stopping time for  $D$ , and  $E_D = \{\tau_D < \infty\}$  is the set of paths that encounter  $D$  after initialization. Let  $\xi$  be a non-trivial,  $\sigma$ -finite measure on  $(\mathcal{W}, \mathcal{D})$  and recall that a set is  $\xi$ -proper if its measure under  $\xi$  is positive and finite.

**Definition 2.1.** A  $\xi$ -proper set  $D$  is *locally  $\xi$ -recurrent* if  $\Pr(E_D|w) = 1$  for all but a  $\xi$ -null set of initial values in  $D$ . Call the Markov chain  $W$  *locally  $\xi$ -recurrent* if every  $\xi$ -proper set is locally  $\xi$ -recurrent.

This notion of recurrence is more general than that typically encountered in general state space Markov chain theory [19], but is appropriate since the chains we will consider in the next section may not be irreducible [10, 14]. We consider a method for establishing local recurrence in Section 3.2.

The connection between the above general Markov chain theory and the notion of strong admissibility relies heavily on the special structure of symmetric Markov chains. Let  $\xi$  be a non-trivial,  $\sigma$ -finite measure. Then the kernel  $K$  is  $\xi$ -symmetric if the measure

$$\omega(A, B) := \int_A K(B|w) \xi(dw)$$

satisfies  $\omega(A, B) = \omega(B, A)$  for all  $A, B \in \mathcal{D}$ . Throughout the remainder we restrict attention to symmetric Markov kernels. Eaton [7, 9] provides some background on the theory of symmetric Markov chains underlying strong admissibility.

### 3. STRONG ADMISSIBILITY VIA MARKOV CHAINS

In a ground-breaking paper Eaton [6] connected the local recurrence of a Markov chain and the almost- $\nu$  admissibility of the formal Bayes estimator for bounded  $\varphi$ . Later, Eaton [8] showed that the local recurrence of a different Markov chain was

required to establish almost- $\nu$  admissibility when  $\varphi$  is unbounded. Since our work builds on them, these foundational definitions and results are stated carefully here. The expected posterior distribution at the parameter value  $\eta$ ,  $R(d\theta|\eta)$  defined in (5), is a  $\nu$ -symmetric Markov kernel on  $\boldsymbol{\vartheta}$  since it satisfies the detailed balance condition

$$(8) \quad R(d\theta|\eta) \nu(d\eta) = R(d\eta|\theta) \nu(d\theta) .$$

Call  $R$  an *Eaton kernel*—analogous kernels exist for the sample space [12] and the product of the sample and parameter spaces [10]. Eaton [6] established the following basic result.

**Theorem 3.1.** *If the Markov chain with kernel  $R$  is locally- $\nu$  recurrent, then the formal Bayes estimator of every bounded function is almost- $\nu$  admissible.*

Eaton [6] also showed that the result holds for bounded perturbations of  $\nu$ . Theorem 3.1 has found substantial application [6, 9, 10, 11, 12, 13, 18].

Now suppose we want to use Markov chains to study formal Bayes estimators of unbounded functions on  $\boldsymbol{\vartheta}$ . We need a basic assumption on the risk function  $R$ , defined at (4), to ensure existence of the integrated risk difference which is studied in the appendices.

**Assumption 3.1.** *Suppose there exist sets  $K_1 \subseteq K_2 \subseteq \dots$  with  $\cup K_i = \boldsymbol{\vartheta}$ . Also assume that for all  $i$ ,  $0 < \nu(K_i) < \infty$  and*

$$\int_{K_i} R(\hat{\varphi}; \theta) \nu(d\theta) < \infty .$$

Let  $\varphi : \boldsymbol{\vartheta} \rightarrow \mathbb{R}^p$  be measurable and unbounded and set

$$\mathcal{S}(\eta) = \int_{\boldsymbol{\vartheta}} \|\varphi(\theta) - \varphi(\eta)\|^2 R(d\theta|\eta)$$

so that if  $0 < \mathcal{S}(\eta) < \infty$  for all  $\eta$ , then

$$(9) \quad S(d\theta|\eta) = \frac{\|\varphi(\theta) - \varphi(\eta)\|^2}{\mathcal{S}(\eta)} R(d\theta|\eta)$$

is a Markov kernel defining a Markov chain on  $\boldsymbol{\vartheta}$ . Moreover  $S$  satisfies detailed balance with respect to  $\mathcal{S}(\eta)\nu(d\eta)$ :

$$S(d\theta|\eta)\mathcal{S}(\eta)\nu(d\eta) = S(d\eta|\theta)\mathcal{S}(\theta)\nu(d\theta) .$$

Eaton [8] connected the local recurrence of  $S$  with the almost- $\nu$  admissibility of the formal Bayes estimator of  $\varphi(\theta)$ .

**Theorem 3.2.** *Suppose Assumption 3.1 holds and assume  $0 < \mathcal{S}(\eta) < \infty$  for each  $\eta \in \mathfrak{V}$  and there exist  $\nu$ -proper sets  $K_1 \subseteq K_2 \subseteq \cdots$  such that  $\cup K_i = \mathfrak{V}$ , and for each  $i$*

$$\int_{K_i} \mathcal{S}(\eta) \nu(d\eta) < \infty .$$

*If the Markov chain with kernel  $S$  is locally- $\nu$  recurrent, then the formal Bayes estimator of  $\varphi$  is almost- $\nu$  admissible.*

In the next section, we unify and generalize these results. We show that by analyzing an appropriate Markov chain we can recover the conclusions of both theorems and, in fact, achieve something stronger. Moreover, we broaden the class of Markov chains that can be studied to obtain strong admissibility results.

**3.1. A new Markov chain connection.** For any  $\eta \in \mathfrak{V}$ , let  $\psi(\cdot|\eta)$  be a nontrivial  $\sigma$ -finite measure on  $(\mathfrak{V}, \mathcal{C})$  such that  $\psi(\cdot|\eta)$  is absolutely continuous with respect to  $R(\cdot|\eta)$ . For any element  $C$  of the Borel sets  $\mathcal{C}$ , let  $\psi(C|\cdot)$  be a nonnegative measurable function. Let  $f(\cdot, \eta)$  be a Radon-Nikodym derivative of  $\psi(\cdot|\eta)$  with respect to  $R(\cdot|\eta)$ —

$$(10) \quad \psi(C|\eta) = \int_C f(\theta, \eta) R(d\theta|\eta) \quad \text{for all } (\eta, C) \in (\mathfrak{V}, \mathcal{C}).$$

Define

$$(11) \quad T(C|\eta) = \frac{\psi(C|\eta)}{\psi(\mathfrak{V}|\eta)} \quad \text{for all } (\eta, C) \in (\mathfrak{V}, \mathcal{C}).$$

Let  $\mathcal{T}(\eta) = \psi(\mathfrak{V}|\eta)$  for all  $\eta \in \mathfrak{V}$ . We will make the following basic assumptions on  $\mathcal{T}$ .

**Assumption 3.2.** *For all  $\eta \in \mathfrak{V}$  we have  $0 < \mathcal{T}(\eta) < \infty$  and there exist  $\nu$ -proper sets  $K_1 \subseteq K_2 \subseteq \cdots$  such that  $\cup K_i = \mathfrak{V}$ , and for each  $i$*

$$\int_{K_i} \mathcal{T}(\eta) \nu(d\eta) < \infty .$$

The first part of the assumption ensures that the kernel  $T$  is well-defined. If  $\mu(d\eta) = \mathcal{T}(\eta) \nu(d\eta)$ , then the second part of the assumption implies that  $\mu$  is  $\sigma$ -finite. In Proposition B.1 we establish that

$$T(d\theta|\eta) \mu(d\eta) = f(\theta, \eta) R(d\theta|\eta) \nu(d\eta) .$$

If  $f(\theta, \eta) = f(\eta, \theta)$  for all  $\eta$  and  $\theta$  in  $\mathfrak{V}$ , then by using (8) it is easy to see that  $T$  is symmetric with respect to  $\mu$ .

Note that  $\mathcal{T}(\eta)$  is almost- $\nu$  uniformly bounded away from 0 if there exists  $\epsilon > 0$  such that  $\mathcal{T}(\eta) \geq \epsilon$  except possibly on a set of  $\nu$ -measure 0. We are now in a position to



state the main result. The proof is given in Appendix B. Recall the definition of  $\Phi_f$  from (6).

**Theorem 3.3.** *Suppose Assumptions 3.1 and 3.2 hold. Let  $\mathcal{T}(\eta)$  be almost- $\nu$  uniformly bounded away from zero and suppose that for all  $\eta$  and  $\theta$  in  $\mathfrak{V}$*

$$(12) \quad f(\theta, \eta) = f(\eta, \theta) .$$

*If the Markov chain having kernel  $T(d\theta|\eta)$  is locally  $\nu$ -recurrent, then  $\nu$  is strongly admissible with respect to  $\Phi_f$ .*

The following extension of Theorem 3.3 is also proved in Appendix B.

**Theorem 3.4.** *Assume the conditions of Theorem 3.3. Then every bounded perturbation  $\nu_u$  is strongly admissible, that is, the family  $\mathcal{F}_\nu$  is strongly admissible with respect to  $\Phi_f$ .*

If we take  $f(\theta, \eta) = d$  for some  $d > 0$ , then we completely recover the results of Theorem 3.1 while if  $f(\theta, \eta) = \|\varphi(\theta) - \varphi(\eta)\|^2$ , then we extend the results of Theorem 3.2. If, as in Examples 1 and 2, we set  $f(\theta, \eta) = \|\varphi(\theta) - \varphi(\eta)\|^2 + d$  with  $d > 0$ , then we obtain results stronger than if we had established local recurrence of the chains associated with the kernels  $R$  and  $S$  and relied on Theorems 3.1 and 3.2. Moreover, since the analyst has the freedom to choose an appropriate  $f$ , this result extends the range of potential applicability of Eaton's method.

**3.2. Reducing dimension.** The Markov kernel  $T$  naturally takes the same dimension as the parameter space. This dimension may be quite large, making the required analysis difficult. In this section we prove that the conclusions of Theorem 3.4 (hence Theorem 3.3) hold if we can establish the local recurrence of a particular Markov chain which lives on  $[0, \infty)$ .

Denote the Borel subsets of  $[0, \infty)$  by  $\mathcal{A}$ . A measurable mapping  $t$  from  $(\mathfrak{V}, \mathcal{C})$  to  $([0, \infty), \mathcal{A})$  induces a measure  $\tilde{\nu}$  given by

$$\tilde{\nu}(A) = \nu(t^{-1}(A)) \quad A \in \mathcal{A} .$$

We will need the next assumption throughout the remainder of this section.

**Assumption 3.3.** *There exists a partition  $\{A_i\}$  of  $[0, \infty)$  such that each  $A_i$  is measurable and each  $C_i = t^{-1}(A_i)$  is  $\nu$ -proper.*

Eaton *et alia* [11] showed that under Assumption 3.3 there exists a Markov transition function  $\pi(d\theta|\beta)$  on  $\mathcal{C} \times [0, \infty)$  such that

$$(13) \quad \nu(d\theta) = \pi(d\theta|\beta) \tilde{\nu}(d\beta)$$

which means for all measurable nonnegative functions  $f_1$  on  $\mathfrak{V}$  and  $f_2$  on  $[0, \infty)$

$$\int_{\mathfrak{V}} f_2(t(\theta)) f_1(\theta) \nu(d\theta) = \int_0^\infty f_2(\beta) \left( \int_{\mathfrak{V}} f_1(\theta) \pi(d\theta|\beta) \right) \tilde{\nu}(d\beta) .$$

Define

$$(14) \quad \tilde{P}(dx|\beta) = \int_{\mathfrak{V}} P(dx|\theta) \pi(d\theta|\beta) \quad \beta \in [0, \infty) .$$

The conditional probabilities  $\tilde{P}(\cdot|\beta)$  form a parametric family indexed by  $\beta$ , and  $\tilde{\nu}(d\beta)$  is a  $\sigma$ -finite prior. Eaton *et alia* [11] also showed that the marginal measure on  $\mathcal{X}$  is the same as at (1). That is,

$$\int_0^\infty \tilde{P}(dx|\beta) \tilde{\nu}(d\beta) = M(dx),$$

which is assumed  $\sigma$ -finite. Thus, there is a Markov kernel  $\tilde{Q}(d\beta|x)$  satisfying

$$(15) \quad \tilde{Q}(d\beta|x) M(dx) = \tilde{P}(dx|\beta) \tilde{\nu}(d\beta) .$$

In fact, a version of the posterior [11] is

$$(16) \quad \tilde{Q}(A|x) := Q(t^{-1}(A)|x) \quad \text{for all } A \in \mathcal{A} .$$

The expected posterior

$$(17) \quad \tilde{R}(d\beta|\alpha) = \int \tilde{Q}(d\beta|x) \tilde{P}(dx|\alpha)$$

is a  $\tilde{\nu}$ -symmetric Eaton kernel.

If  $\eta$  and  $\theta$  are elements of  $\mathfrak{V}$ , let  $\alpha = t(\eta)$  and  $\beta = t(\theta)$ . For any  $\alpha \in [0, \infty)$ , let  $\tilde{\psi}(\cdot|\alpha)$  be a nontrivial  $\sigma$ -finite measure on  $([0, \infty), \mathcal{A})$  such that  $\tilde{\psi}(\cdot|\alpha)$  is absolutely continuous with respect to  $\tilde{R}(\cdot|\alpha)$ . Let  $\tilde{f}(\cdot, \alpha)$  be a Radon-Nikodym derivative of  $\tilde{\psi}(\cdot|\alpha)$  with respect to  $\tilde{R}(\cdot|\alpha)$ —ie, for all  $(\alpha, A) \in ([0, \infty), \mathcal{A})$

$$\tilde{\psi}(A|\alpha) = \int_A \tilde{f}(\beta, \alpha) \tilde{R}(d\beta|\alpha) .$$

Define  $\tilde{\mathcal{T}}(\alpha) = \tilde{\psi}([0, \infty)|\alpha)$  and set

$$\tilde{T}(A|\alpha) = \frac{\tilde{\psi}(A|\alpha)}{\tilde{\mathcal{T}}(\alpha)} \quad (\alpha, A) \in ([0, \infty), \mathcal{A})$$

and let  $\tilde{\mu}(d\alpha) = \tilde{T}(\alpha) \tilde{\nu}(d\alpha)$ . The following assumption ensures that  $\tilde{T}$  is a well-defined kernel and that  $\tilde{\mu}$  is  $\sigma$ -finite.

**Assumption 3.4.** *For all  $\alpha$  we have  $0 < \tilde{T}(\alpha) < \infty$  and there exist  $\tilde{\nu}$ -proper sets  $K_1 \subseteq K_2 \subseteq \dots$  such that  $\cup K_i = [0, \infty)$ , and for each  $i$*

$$\int_{K_i} \tilde{T}(\alpha) \tilde{\nu}(d\alpha) < \infty .$$

In Appendix C we show that  $\tilde{T}$  is  $\tilde{\mu}$ -symmetric. The following theorem shows that we can analyze the recurrence properties of the chain defined by  $\tilde{T}$  to achieve the conclusions of Theorem 3.4 (hence Theorem 3.3). The proof is given in Appendix C while use of the result is illustrated in Section 4. Recall that  $\mathcal{F}_\nu$  is the family of bounded perturbations of  $\nu$ .

**Theorem 3.5.** *Suppose Assumption 3.4 holds. Let  $\tilde{T}(\cdot)$  be almost- $\tilde{\nu}$  uniformly bounded away from 0 and suppose that for all  $\alpha$  and  $\beta$  in  $[0, \infty)$*

$$\tilde{f}(\beta, \alpha) = \tilde{f}(\alpha, \beta) .$$

*If the Markov chain with kernel  $\tilde{T}$  is locally  $\tilde{\nu}$ -recurrent, then  $\mathcal{F}_\nu$  is strongly admissible with respect to  $\Phi_f$ .*

Since  $\tilde{T}$  lives on  $[0, \infty)$  it would be convenient to have conditions which guarantee the local recurrence of a Markov chain on  $[0, \infty)$ . This is discussed in the following section.

**3.2.1. Recurrence of Markov chains on  $[0, \infty)$ .** To this point we have said little about establishing local recurrence. The following theorem presents one method for doing so and is a distillation of several existing results [9, 11]. It applies generally to Markov chains on  $[0, \infty)$  and hence the notation in this section is consistent with that of Section 2.

Let  $\mathcal{W} = [0, \infty)$  and  $\mathcal{D}$  be the Borel  $\sigma$ -algebra. Let  $K : \mathcal{D} \times \mathcal{W} \rightarrow [0, 1]$  be a Markov kernel which defines a time-homogeneous Markov chain  $W = \{W_0, W_1, W_2, \dots\}$  on  $\mathcal{W}^\infty$ . Define the  $k$ th moment of  $K$  about its current state as

$$(18) \quad m_k(v) = \int_0^\infty (x - v)^k K(dx|v) .$$

**Theorem 3.6.** *Assume for each positive integer  $n$  there exists  $\delta(n) < 1$  such that*

$$(19) \quad \sup_{v \in [0, n]} K([0, n]|v) \leq \delta(n) .$$

Suppose

$$(20) \quad \lim_{v \rightarrow \infty} \frac{\log v}{v} \frac{m_3(v)}{m_2(v)} = 0$$

and there exists a function  $\phi$  and an integer  $n_0$  such that for  $v \in [n_0, \infty)$

$$(21) \quad m_1(v) \leq \frac{m_2(v)}{2v} [1 + \phi(v)] \quad \text{and} \quad \lim_{v \rightarrow \infty} \phi(v) \log v = 0.$$

If  $\xi$  is a non-trivial,  $\sigma$ -finite measure,  $0 < \xi([0, n_0)) < \infty$  and  $K$  is  $\xi$ -symmetric, then the Markov chain  $W$  is locally  $\xi$ -recurrent.

We will use Theorem 3.6 in conjunction with Theorem 3.5 in our main application in Section 4.

#### 4. ADMISSIBLE PRIORS FOR THE MULTIVARIATE NORMAL MEAN

Let  $X \sim N_p(\theta, I_p)$  with  $p \geq 1$ . Consider the family of  $\sigma$ -finite measures on  $\boldsymbol{\vartheta} = \mathbb{R}^p$  described by

$$(22) \quad \left( \frac{1}{a + \|\theta\|^2} \right)^b d\theta \quad a \geq 0, \quad b \geq 0.$$

For  $a = 0$ , the prior is improper for all  $b > 0$ , but the induced marginal distributions on  $\mathcal{X}$  are only  $\sigma$ -finite for  $b < p/2$ . For  $a > 0$ , the family yields improper prior distributions when  $b \leq p/2$  and proper prior distributions when  $b > p/2$ . In fact, if  $b = (a + p)/2$ , the prior is the kernel of a multivariate  $t$  distribution with  $a$  degrees of freedom. When  $a > 0$  and  $b = 0$ , this is  $p$ -dimensional Lebesgue measure which was considered in Example 1.

Now suppose  $p \geq 3$ . Berger, Strawderman and Tang [3] established that the formal Bayes estimator of  $\theta$  is admissible when  $a = 1$  and  $b = (p - 1)/2$  while Eaton *et alia* [11] use Theorem 3.1 to prove that if  $a > 0$  and  $b \in [p/2 - 1, p/2]$ , then the formal Bayes estimator of every bounded function is almost admissible. We use Theorems 3.5 and 3.6 to add to these results. Let  $d > 0$  be arbitrary and define

$$f(\theta, \eta) = \|\theta - \eta\|^2 + d \quad \theta, \eta \in \mathbb{R}^p$$

so that

$$\Phi_f = \{\varphi : \|\varphi(\theta) - \varphi(\eta)\| \leq M_\varphi f(\theta, \eta) \quad \text{some } 0 < M_\varphi < \infty\}.$$

**Theorem 4.1.** For  $p \geq 3$  let  $X \sim N_p(\theta, I_p)$  and with  $a > 0$  set

$$\nu(d\theta) = \left( \frac{1}{a + \|\theta\|^2} \right)^{p/2} d\theta.$$

Then  $\mathcal{F}_\nu$ , the family of bounded perturbations of  $\nu$ , is strongly admissible with respect to  $\Phi_f$ .

*Proof of Theorem 4.1.* Let  $\gamma_a(z) = (a + z)^{-p/2}$ , so that our family of priors (22) can be expressed as  $\gamma_a(\|\theta\|^2) d\theta$ . The function  $t(\theta) = \|\theta\|^2$  fulfills the requirements of Assumption 3.3 as can be seen by letting the sets  $A_i = [i-1, i)$  partition the nonnegative real numbers. Letting  $\pi(d\theta|\beta)$  denote the uniform distribution on the hypersphere of radius  $\sqrt{\beta}$ , it can be shown that  $\nu(d\theta) = \pi(d\theta|\beta) \tilde{\nu}(d\beta)$  with

$$\tilde{\nu}(d\beta) = \frac{[\Gamma(1/2)]^p}{\Gamma(p/2)} \gamma_a(\beta) \beta^{\frac{p}{2}-1} d\beta$$

on  $[0, \infty)$ . Let  $\pi_1$  be the uniform distribution on the unit hypersphere  $\Xi$ . Then the reduced sampling distribution has density

$$\tilde{p}(x|\beta) = \int_{\Xi} (2\pi)^{-p/2} \exp\left(-\frac{1}{2}\|x - \xi\sqrt{\beta}\|^2\right) \pi_1(d\xi)$$

with respect to Lebesgue measure on  $\mathcal{X}$ . If

$$m(x) = \int_0^\infty \tilde{p}(x|\beta) \tilde{\nu}(d\beta),$$

then the formal posterior has density

$$\tilde{q}(\beta|x) = \frac{[\Gamma(1/2)]^p}{\Gamma(p/2)} \frac{\tilde{p}(x|\beta) \gamma_a(\beta) \beta^{\frac{p}{2}-1}}{m(x)}$$

with respect to Lebesgue measure on  $[0, \infty)$ . The expected posterior

$$(23) \quad \tilde{R}(d\beta|\alpha) = \int_{\mathcal{X}} \tilde{q}(\beta|x) \tilde{p}(x|\alpha) dx d\beta$$

is a  $\tilde{\nu}$ -symmetric Markov transition kernel on  $[0, \infty)$ .

Let  $\tilde{f}(\beta, \alpha) = 2(\beta + \alpha + c)$  with  $c$  a positive constant. It is clear that  $\tilde{f}$  is symmetric in  $\alpha$  and  $\beta$  and bounded away from zero. For any nonnegative real number  $\alpha$ ,

$$\tilde{\mathcal{T}}(\alpha) = \int \tilde{f}(\beta, \alpha) \tilde{R}(d\beta|\alpha) = 2 \int \beta \tilde{R}(d\beta|\alpha) + 2\alpha + 2c$$

is greater than or equal to  $c$ . Thus the kernel

$$(24) \quad \tilde{T}(d\beta|\alpha) = \frac{(\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}$$

defines a  $\tilde{\mu}$ -symmetric Markov chain on  $[0, \infty)$  where

$$\tilde{\mu}(A) = \int_A \tilde{T}(\alpha) \tilde{\nu}(d\alpha) \quad A \in \mathcal{A}.$$

Clearly, for any integer  $n$ ,  $\tilde{\mu}([0, n]) < \infty$ .

The next step is to verify the conditions of Theorem 3.6 which will imply the Markov chain associated with  $\tilde{T}$  is locally  $\tilde{\mu}$ -recurrent. Since  $\tilde{\mu}$  and  $\tilde{\nu}$  are equivalent measures we will also conclude that the chain is locally  $\tilde{\nu}$ -recurrent. In Appendix D it is shown that  $\tilde{T}([0, m]|\alpha)$  is continuous as a function of  $\alpha$  implying condition (19) of Theorem 3.6. For  $k$  a nonnegative integer set

$$m_k(\alpha) = \int (\beta - \alpha)^k \tilde{T}(d\beta|\alpha).$$

Additional calculations given in Appendix D show that

$$(25) \quad \frac{m_3(\alpha)}{m_2(\alpha)} = O(1) \quad \text{as } \alpha \rightarrow \infty$$

and

$$(26) \quad m_1(\alpha) = \frac{8\alpha + \psi_1(\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}$$

where  $\psi_1(\alpha) = O(1)$  as  $\alpha \rightarrow \infty$ . Letting  $\phi(\alpha) = 1/\sqrt{\alpha}$ , note that  $\lim_{\alpha \rightarrow \infty} \phi(\alpha) \log(\alpha) = 0$  and, by calculations in Appendix D,

$$(27) \quad \frac{m_2(\alpha)}{2\alpha} [1 + \phi(\alpha)] = \frac{8\alpha + 8\sqrt{\alpha} + \psi_2(\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}$$

where  $\psi_2(\alpha) = O(1)$  as  $\alpha \rightarrow \infty$ . It is clear by inspection of (25), (26), and (27) that the conditions (20) and (21) of Theorem 3.6 are satisfied for  $\alpha$  large enough. Hence the chain is locally  $\tilde{\nu}$ -recurrent.

By Theorem 3.5, the family of priors  $\mathcal{F}_\nu$  is strongly admissible with respect to  $\Phi_f$ .

□

## APPENDIX A. PRELIMINARIES

We begin by stating some existing results concerning local recurrence and introduce Blyth's method. This material plays a fundamental role in our proofs of Theorems 3.3, 3.4 and 3.5.

**A.1. Local Recurrence.** The purpose here is to give two characterizations of local recurrence for general symmetric Markov chains, hence the notation is consistent with that of Sections 2 and 3.2.1.

Let  $\mathcal{W}$  be a Polish space and denote the Borel  $\sigma$ -algebra by  $\mathcal{D}$ . Let  $K : \mathcal{D} \times \mathcal{W} \rightarrow [0, 1]$  be a Markov kernel on  $(\mathcal{W}, \mathcal{D})$ . Let  $\xi$  be a non-trivial  $\sigma$ -finite measure and recall that a set is  $\xi$ -proper if its measure under  $\xi$  is positive and finite. Throughout this section  $K$  is assumed to be  $\xi$ -symmetric.

**Theorem A.1** (Eaton [9]). *The Markov chain  $W$  is locally  $\xi$ -recurrent if and only if there exists a sequence of  $\xi$ -proper sets increasing to the state space such that each is locally  $\xi$ -recurrent.*

Let  $L^2(\xi)$  be the space of  $\xi$ -square integrable functions. Then the quantity

$$(28) \quad \Delta(h; K, \xi) = \frac{1}{2} \iint (h(\theta) - h(\eta))^2 K(d\theta|\eta) \xi(d\eta) \quad h \in L^2(\xi)$$

is called a *Dirichlet form*. Also, if  $D$  is a  $\xi$ -proper set, define

$$\mathcal{H}_\xi(D) := \{h \in L^2(\xi) : h \geq I_D\}$$

where  $I_D$  is the indicator function of the set  $D$ . A characterization of local  $\xi$ -recurrence in terms of  $\Delta$  is given by the following result.

**Theorem A.2** (Eaton [8]). *A set  $D \in \mathcal{D}$  is locally  $\xi$ -recurrent if and only if*

$$\inf_{\mathcal{H}_\xi(D)} \Delta(h; K, \xi) = 0 .$$

**A.2. Blyth's Method.** Consider the posterior distributions obtained from perturbations of the prior measure  $\nu$ . Let  $g : \mathfrak{V} \rightarrow \mathbb{R}^+$  be such that the perturbation

$$(29) \quad \nu_g(d\theta) = g(\theta)\nu(d\theta)$$

is  $\sigma$ -finite. Assume

$$(30) \quad \hat{g}(x) := \int g(\theta)Q(d\theta|x) < \infty .$$

Letting

$$M_g(dx) := \int P(dx|\theta)g(\theta)\nu(d\theta)$$

it is easy to see that  $M_g(dx) = \hat{g}(x)M(dx)$  is  $\sigma$ -finite and that the posterior obtained from the perturbed prior is

$$(31) \quad Q_g(d\theta|x) = \frac{g(\theta)}{\hat{g}(x)}Q(d\theta|x) .$$

Let  $L^1(\nu)$  be the set of all  $\nu$ -integrable functions and define

$$(32) \quad \mathcal{G}_\nu = \{g \in L^1(\nu) : g \geq 0, g \text{ bounded and } \nu_g(\boldsymbol{\vartheta}) > 0\}.$$

Let  $I_C$  be the indicator function of the set  $C$ . Of particular interest are the subfamilies

$$\mathcal{G}_\nu(C) = \{g \in \mathcal{G}_\nu : g \geq I_C\}.$$

Let  $\hat{\varphi}$  be the Bayes estimator of  $\varphi$  under the prior  $\nu$ , and let  $\hat{\varphi}_g$  be the Bayes estimator of  $\varphi$  under the prior  $\nu_g$ . Also recall the definition of the risk function  $R$  at (4). A key quantity in connecting Markov chains to admissibility is the *integrated risk difference* at  $g \in \mathcal{G}_\nu$  with respect to  $\nu$  against  $\hat{\varphi}$ :

$$(33) \quad \text{IRD}_\nu(g; \hat{\varphi}) := \int [\mathbf{R}(\theta; \hat{\varphi}) - \mathbf{R}(\theta; \hat{\varphi}_g)] \nu_g(d\theta).$$

Notice that the integrability assumptions on  $R$  in Assumption 3.1 ensure that  $\text{IRD}_\nu$  is well-defined. The connection of  $\text{IRD}_\nu$  with almost- $\nu$  admissibility is given by Blyth's method.

**Theorem A.3** (Blyth's Method). *Let  $\hat{\varphi}$  be an estimator and  $\nu$  a  $\sigma$ -finite measure on the parameter space. If  $\inf_{g \in \mathcal{G}_\nu(C)} \text{IRD}_\nu(g; \hat{\varphi}) = 0$  for every  $C \subseteq \boldsymbol{\vartheta}$  such that  $0 < \nu(C) < \infty$ , then  $\hat{\varphi}$  is almost- $\nu$  admissible.*

## APPENDIX B. PROOF OF THEOREM 3.3

We will develop a connection between the kernel  $T$  defined in (11) and the  $\text{IRD}_\nu$  which will be key to proving Theorem 3.3 via Theorems A.2 and A.3. We begin with some preliminary results before we prove Theorem 3.3.

**B.1. Preliminary Results.** Recall  $\mathcal{T}(\eta) = \psi(\boldsymbol{\vartheta}|\eta)$  where  $\psi$  is defined at (10). Define

$$(34) \quad \mu(C) = \int_C \mathcal{T}(\eta) \nu(d\eta) \quad \text{for all } C \in \mathcal{C}.$$

**Proposition B.1.** *Suppose Assumption 3.2 and condition (12) hold. Then*

- (a) *the measure  $\mu$  is  $\sigma$ -finite and equivalent to  $\nu$ ,*
- (b)  *$\mathcal{T}$  is a Radon-Nikodym derivative of  $\mu$  with respect to  $\nu$ ,*
- (c)  *$T$  is a  $\mu$ -symmetric Markov transition kernel,*
- (d)  *$T(d\theta|\eta)\mu(d\eta) = f(\theta, \eta)R(d\theta|\eta)\nu(d\eta)$ ,*
- (e) *if  $\mu'$  is proportional to  $\mu$ ,  $T$  is a  $\mu'$ -symmetric Markov transition kernel.*



*Proof.* That  $\mu$  is  $\sigma$ -finite follows easily from Assumption 3.2. Since  $\mathcal{T}$  is a nonnegative measurable function, for any  $\nu$ -null set  $C$ ,

$$\mu(C) = \int_C \mathcal{T}(\eta) \nu(d\eta) = 0 .$$

Furthermore, since  $\mathcal{T}(\eta) > 0$  for all  $\eta$  in the parameter space, every  $\nu$ -positive set is a  $\mu$ -positive set. Thus, if  $C$  is  $\mu$ -null, it is also  $\nu$ -null. Therefore,  $\mu$  and  $\nu$  are equivalent measures. Furthermore,  $\mathcal{T}$  is a Radon-Nikodym derivative of  $\mu$  with respect to  $\nu$  since  $\mu$  is absolutely continuous with respect to  $\nu$ , and by assumption  $\mathcal{T}$  is a nonnegative measurable function such that for any measurable set  $C$ ,  $\mu(C)$  is given by equation (34).

At any point  $\eta$  in the parameter space, since  $\psi(\cdot|\eta)$  is a nontrivial finite measure, normalizing by  $\mathcal{T}(\eta) = \psi(\boldsymbol{\vartheta}|\eta)$  produces a probability measure. Recall that  $\psi(C|\cdot)$  is a nonnegative measurable function where  $C$  is an element of the Borel sets  $\mathcal{C}$ . Since  $\boldsymbol{\vartheta}$  is a Borel measurable set,  $\mathcal{T}$  is a Borel measurable function. Also by hypothesis,  $\mathcal{T}(\eta)$  is positive and finite for all  $\eta$ . The reciprocal function is continuous, hence Borel measurable, on the positive real numbers, and the composition  $1/\mathcal{T}$  of Borel measurable functions is a Borel measurable function. Thus, the product of Borel measurable functions  $1/\mathcal{T}$  and  $\psi(C|\cdot)$  is a Borel measurable function for any  $C \in \mathcal{C}$ .

Let  $A$  and  $B$  be measurable sets. Since

$$\begin{aligned} S(A, B) &= \iint I_A(\eta) I_B(\theta) T(d\theta|\eta) \mu(d\eta) \\ &= \iint I_A(\eta) I_B(\theta) f(\theta, \eta) R(d\theta|\eta) \nu(d\eta) && \text{by (10), (11), and (34)} \\ &= \iint I_B(\theta) I_A(\eta) f(\eta, \theta) R(d\eta|\theta) \nu(d\theta) && \text{by (8) and (12)} \\ &= \iint I_B(\theta) I_A(\eta) T(d\eta|\theta) \mu(d\theta) && \text{by (10), (11), and (34)} \\ &= S(B, A), \end{aligned}$$

the kernel  $T$  is  $\mu$ -symmetric. Note that the substitution

$$(35) \quad T(d\theta|\eta) \mu(d\eta) = f(\theta, \eta) R(d\theta|\eta) \nu(d\eta)$$

holds as a consequence.

Let  $\mu' = c\mu$  for some  $c > 0$ . Then by (8) and (35)

$$\begin{aligned}
T(d\theta|\eta)\mu'(d\eta) &= cT(d\theta|\eta)\mu(d\eta) \\
&= cf(\theta, \eta)R(d\theta|\eta)\nu(d\eta) \\
&= cf(\eta, \theta)R(d\eta|\theta)\nu(d\theta) \\
&= cT(d\eta|\theta)\mu(d\theta) \\
&= T(d\eta|\theta)\mu'(d\theta)
\end{aligned}$$

which implies  $T$  is  $\mu'$ -symmetric.  $\square$

We can now develop a connection between the Markov kernel  $T$  and the integrated risk difference  $\text{IRD}_\nu$ . Our argument will require the following known result; recall the definition of  $\mathcal{G}_\nu$  from (32).

**Proposition B.2** (Eaton, 2001). *If  $g \in \mathcal{G}_\nu$  and  $R$  is Eaton's kernel, then*

$$\text{IRD}_\nu(g; \hat{\varphi}) \leq \iint \|\varphi(\theta) - \varphi(\eta)\|^2 (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 R(d\theta|\eta) \nu(d\eta) .$$

By Proposition B.1 we see that  $T$  is  $\mu$ -symmetric so that the relevant Dirichlet form, recall (28), is

$$\Delta(h; T, \mu) = \frac{1}{2} \iint (h(\eta) - h(\theta))^2 T(d\theta|\eta) \mu(d\eta) \quad h \in L^2(\mu) .$$

**Proposition B.3.** *Suppose Assumption 3.2 and condition (12) hold. If  $\varphi \in \Phi_f$  and  $g \in \mathcal{G}_\nu$ , then there is a measure  $\mu_\varphi$  which is proportional to  $\mu$  such that*

$$\text{IRD}_\nu(g; \hat{\varphi}) \leq 2\Delta(\sqrt{g}; T, \mu_\varphi) .$$

*Proof.* Since  $\varphi \in \Phi_f$  there exists  $0 < M_\varphi < \infty$  such that

$$\|\varphi(\theta) - \varphi(\eta)\|^2 \leq M_\varphi f(\theta, \eta) \quad \text{for all } \theta, \eta$$

For  $C \in \mathcal{C}$  define  $\mu_\varphi(C) = M_\varphi \mu(C)$  and suppose  $g \in \mathcal{G}_\nu$ . In the following the first inequality is from Proposition B.2 and the second is obtained from Proposition B.1(d) and that  $\varphi \in \Phi_f$

$$\begin{aligned}
\text{IRD}_\nu(g; \hat{\varphi}) &\leq \iint \|\varphi(\theta) - \varphi(\eta)\|^2 (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 R(d\theta|\eta) \nu(d\eta) \\
&\leq \iint (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 T(d\theta|\eta) M_\varphi \mu(d\eta) \\
&= 2\Delta(\sqrt{g}; T, \mu_\varphi) .
\end{aligned}$$

□

**B.2. Proof of Theorem 3.3.** Let the measures  $\psi(\cdot|\eta)$  and  $\mu$  and the transition kernel  $T$  be as defined at equations (10), (34), and (11), respectively. Also, suppose  $\varphi \in \Phi_f$ . By Proposition B.3 if  $g \in \mathcal{G}_\nu$ , then

$$\text{IRD}_\nu(g; \hat{\varphi}) \leq 2\Delta(\sqrt{g}; T, \mu_\varphi)$$

where  $\mu_\varphi = M_\varphi \mu$ . By Proposition B.1 we have that  $\mu$  and  $\nu$  are equivalent measures and  $T$  is a  $\mu_\varphi$ -symmetric Markov kernel. Since the measures  $\mu_\varphi$  and  $\nu$  are equivalent and the chain with kernel  $T$  is locally  $\nu$ -recurrent, it is also locally  $\mu_\varphi$ -recurrent—that is, every  $\mu_\varphi$ -proper set  $C$  is locally  $\mu_\varphi$ -recurrent. Thus, by Theorem A.2,

$$\inf_{\mathcal{H}_{\mu_\varphi}(C)} \Delta(h; T, \mu_\varphi) = 0$$

where  $\mathcal{H}_{\mu_\varphi}(C)$  collects the square-integrable dominators of  $I_C$ .

Now let  $h \in \mathcal{H}_{\mu_\varphi}(C)$  and recall that by assumption  $\mathcal{T}$  is a measurable function uniformly bounded away from zero  $\nu$ -almost everywhere. Hence there exists some  $\epsilon > 0$  such that

$$\begin{aligned} \epsilon \int h^2(\eta) \nu(d\eta) &\leq \int h^2(\eta) \mathcal{T}(\eta) \nu(d\eta) \\ &= \int h^2(\eta) \mu(d\eta) \\ &= \frac{1}{M_\varphi} \int h^2(\eta) \mu_\varphi(d\eta) \\ &< \infty. \end{aligned}$$

Hence  $h \in L^2(\nu)$  and we conclude that  $\mathcal{H}_{\mu_\varphi}(C) \subseteq \mathcal{G}_\nu(C)$ . Moreover, if  $\sqrt{h} \in \mathcal{H}_{\mu_\varphi}(C)$ , then  $h \in \mathcal{G}_\nu(C)$ . Thus we obtain

$$\inf_{\mathcal{G}_\nu(C)} \text{IRD}_\nu(g; \hat{\varphi}) = 0.$$

Therefore,  $\hat{\varphi}$  is an almost- $\nu$  admissible estimator by Theorem A.3.

**B.3. Proof of Theorem 3.4.** Recall that  $\nu_u \in \mathcal{F}_\nu$  is a bounded perturbation of  $\nu$ . The sampling and posterior distributions define the  $\nu_u$ -symmetric Eaton kernel

$$R_u(d\theta|\eta) = \int_{\mathcal{X}} Q_u(d\theta|x) P(dx|\eta)$$

where the perturbed posterior was defined at (31). The mean of  $\varphi$  with respect to  $Q_u(\cdot|x)$  is the formal Bayes estimator of  $\varphi(\theta)$  under squared error loss. We denote it  $\hat{\varphi}_u$  to emphasize its dependence on the perturbed prior.

Since  $\nu_u \in \mathcal{F}_\nu$ , there exists  $0 < c < \infty$  such that  $\nu_u(d\theta) = u(\theta)\nu(d\theta)$  and  $1/c < u(\theta) < c$ . Recall the definition of  $\mu_\varphi$  from the proof of Proposition B.3. In the following the first inequality is from Proposition B.2, the second follows by noting that  $u$  and  $1/\hat{u}$  are both bounded above by  $c$  while the third is obtained from Proposition B.1(d) and that  $\varphi \in \Phi_f$

$$\begin{aligned} \text{IRD}_{\nu_u}(g; \hat{\varphi}_u) &\leq \iint \|\varphi(\theta) - \varphi(\eta)\|^2 (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 R_u(d\theta|\eta) \nu_u(d\eta) \\ &\leq c^3 \iint \|\varphi(\theta) - \varphi(\eta)\|^2 (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 R(d\theta|\eta) \nu(d\eta) \\ &\leq c^3 \iint (\sqrt{g}(\theta) - \sqrt{g}(\eta))^2 T(d\theta|\eta) M_\varphi \mu(d\eta) \\ &= 2c^3 \Delta(\sqrt{g}; T, \mu_\varphi). \end{aligned}$$

The remainder of the proof follows the proof of Theorem 3.3 exactly.

### APPENDIX C. PROOF OF THEOREM 3.5

Before proving Theorem 3.5 we require an analogue of Proposition B.1. Given a nonnegative function  $\tilde{h}$  on  $[0, \infty)$ , define

$$(36) \quad h(\theta) = \tilde{h}(t(\theta)).$$

**Proposition C.1.** *Suppose Assumptions 3.3 and 3.4 hold. If  $\tilde{f}(\beta, \alpha) = \tilde{f}(\alpha, \beta)$ , then*

- (a)  $\tilde{\mu}$  is  $\sigma$ -finite and equivalent to  $\tilde{\nu}$ ,
- (b)  $\tilde{T}$  is a Radon-Nikodym derivative of  $\tilde{\mu}$  with respect to  $\tilde{\nu}$ ,
- (c)  $\tilde{T}$  is a  $\tilde{\mu}$ -symmetric Markov transition kernel,
- (d)  $\tilde{T}(d\beta|\alpha) \tilde{\mu}(d\alpha) = \tilde{f}(\beta, \alpha) \tilde{R}(d\beta|\alpha) \tilde{\nu}(d\alpha)$ ,
- (e)  $\tilde{h} \in L^2(\tilde{\mu})$  implies  $h$  is in  $L^2(\mu)$ ,
- (f)  $\Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = \Delta(h; T, \mu)$ , and
- (g)  $\mathcal{T}(\eta)$  is almost- $\nu$  bounded away from zero.

*Proof.* The proof of the first 3 assertions follows exactly the proof of the first 3 assertions in Proposition B.1 with  $[0, \infty)$ ,  $\mathcal{A}$ ,  $\alpha$ ,  $\beta$ ,  $\tilde{\nu}$ , and  $\tilde{R}$  substituted for  $\boldsymbol{\vartheta}$ ,  $\mathcal{C}$ ,  $\eta$ ,  $\theta$ ,  $\nu$ , and  $R$ , respectively. The substitution

$$(37) \quad \tilde{T}(d\beta|\alpha) \tilde{\mu}(d\alpha) = \tilde{f}(\beta, \alpha) \tilde{R}(d\beta|\alpha) \tilde{\nu}(d\alpha)$$

follows as a consequence. If  $\tilde{h} \in L^2(\tilde{\mu})$ , then  $h \in L^2(\mu)$  since

$$\begin{aligned}
\int \tilde{h}^2(\alpha) \tilde{\mu}(d\alpha) &= \int \tilde{h}^2(\alpha) \tilde{\mathcal{T}}(\alpha) \tilde{\nu}(d\alpha) \\
&= \int \tilde{h}^2(\alpha) \int \tilde{f}(\beta, \alpha) \int \tilde{Q}(d\beta|x) \tilde{P}(dx|\alpha) \tilde{\nu}(d\alpha) \\
&= \int h^2(\eta) \int f(t(\theta), t(\eta)) \int Q(d\theta|x) P(dx|\eta) \nu(d\eta) \\
&= \int h^2(\eta) \mathcal{T}(\eta) \nu(d\eta) \\
&= \int h^2(\eta) \mu(d\eta).
\end{aligned}$$

Since

$$\begin{aligned}
&\iint (\tilde{h}(\beta) - \tilde{h}(\alpha))^2 \tilde{T}(d\beta|\alpha) \tilde{\mu}(d\alpha) \\
&= \iint (\tilde{h}(\beta) - \tilde{h}(\alpha))^2 \tilde{f}(\beta, \alpha) \tilde{R}(d\beta|\alpha) \tilde{\nu}(d\alpha) && \text{by (37)} \\
&= \iint (\tilde{h}(\beta) - \tilde{h}(\alpha))^2 \tilde{f}(\beta, \alpha) \int \tilde{Q}(d\beta|x) \tilde{P}(dx|\alpha) \tilde{\nu}(d\alpha) && \text{by (17)} \\
&= \iint (h(\theta) - h(\eta))^2 f(t(\theta), t(\eta)) \int Q(d\theta|x) P(dx|\eta) \nu(d\eta) && \text{by (13), (14), (16), and (36)} \\
&= \iint (h(\theta) - h(\eta))^2 f(t(\theta), t(\eta)) R(d\theta|\eta) \nu(d\eta) && \text{by (5)} \\
&= \iint (h(\theta) - h(\eta))^2 T(d\theta|\eta) \mu(d\eta) && \text{by (35),}
\end{aligned}$$

it follows that  $\Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = \Delta(h; T, \mu)$ . Finally, recall that  $\tilde{\mathcal{T}}$  is almost- $\tilde{\nu}$  bounded away from zero by some positive constant  $c$ . Note that for any  $\tilde{\nu}$ -proper set  $A$ ,

$$\int_{t^{-1}(A)} \mathcal{T}(\eta) \nu(d\eta) = \int_A \tilde{\mathcal{T}}(\alpha) \tilde{\nu}(d\alpha) \geq c \tilde{\nu}(A) = c \nu(t^{-1}(A)).$$

It follows that  $\mathcal{T}$  is almost- $\nu$  bounded away from zero.  $\square$

We are now ready to prove Theorem 3.5.

**C.1. Proof of Theorem 3.5.** By Proposition C.1, the measures  $\tilde{\mu}$  and  $\tilde{\nu}$  are equivalent, and the kernel  $\tilde{T}$  is  $\tilde{\mu}$ -symmetric. Thus, the chain with kernel  $\tilde{T}$  is locally  $\tilde{\mu}$ -recurrent, and by Theorem A.2

$$\inf_{\mathcal{H}_{\tilde{\mu}}(A)} \Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = 0$$

for any  $\tilde{\mu}$ -proper set  $A$ .

Let  $(A_i)$  be a sequence of  $\tilde{\mu}$ -proper sets increasing to  $[0, \infty)$ . Letting  $C_i = t^{-1}(A_i)$  defines a sequence of  $\mu$ -proper sets increasing to  $\mathfrak{I}$ . Since  $h(\theta) = \tilde{h}(t(\theta))$ ,  $\tilde{h} \geq I_{A_i}$  implies that  $h \geq I_{C_i}$ . By Proposition C.1,  $\tilde{h} \in L^2(\tilde{\mu})$  implies  $h \in L^2(\mu)$  and the corresponding Dirichlet forms are equal. Thus,

$$\inf_{\mathcal{H}_\mu(C_i)} \Delta(h; T, \mu) = \inf_{\mathcal{H}_{\tilde{\mu}}(A_i)} \Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = 0$$

for every  $C_i$  in the sequence. By Theorem A.1, the chain with kernel  $T$  is locally  $\mu$ -recurrent. Since  $\mu$  and  $\nu$  are equivalent measures, the chain is locally  $\nu$ -recurrent as well.

By Proposition C.1,  $\mathcal{T}$  is almost- $\nu$  bounded away from zero, and thus all of the conditions for Theorem 3.4 are satisfied. Therefore, under squared error loss, the family  $\mathcal{F}_\nu$  is  $\Phi_f$ -admissible.

#### APPENDIX D. MULTIVARIATE NORMAL

Here we give the supplemental arguments required for the proof of Theorem 4.1.

**D.1. Existence of the integrated risk difference.** In order to appeal to Blyth's method, we need to know that the integrated risk differences are defined. Otherwise the bounding inequality (B.2) for the integrated risk differences is meaningless. The proper Bayes estimators necessarily have finite integrated risks, so it is sufficient to show that

$$\int \mathbf{R}(\theta; \hat{\theta}) g_n(\theta) \nu(d\theta) < \infty$$

where  $g_n$  is the indicator of  $C_n$  a closed ball around zero with radius  $n$ . Since  $\nu$  is  $\sigma$ -finite,  $\nu(C_n)$  is finite. Since the risk function is real-valued and continuous, it attains a finite maximum on  $C_n$ ; call it  $M_n$ . Therefore,

$$\int \mathbf{R}(\theta; \hat{\theta}) g_n(\theta) \nu(d\theta) \leq M_n \nu(C_n) < \infty$$

and the integrated risk differences are defined.

**D.2. Continuity of the transition kernel.** Let  $m$  be a nonnegative integer. We wish to show that  $T([0, m]|\alpha)$  is continuous as a function of  $\alpha$ . Since by definition

$$T([0, m]|\alpha) = \frac{\int_0^m \beta \tilde{R}(d\beta|\alpha) + \alpha \tilde{R}([0, m]|\alpha) + c \tilde{R}([0, m]|\alpha)}{\int \beta \tilde{R}(d\beta|\alpha) + \alpha + c}$$

it is sufficient to show that  $\tilde{R}([0, m]|\alpha)$  is continuous and that continuity of

$$\int_0^m \beta \tilde{R}(d\beta|\alpha) \quad \text{and} \quad \int \beta \tilde{R}(d\beta|\alpha)$$

follows from there.

Fix  $\alpha^* \in [0, \infty)$  and  $\delta$  greater than zero. Let  $S_\delta(\alpha^*) = [0, (\sqrt{\alpha^*} + \delta)^2)$ . Let  $(\alpha_n)$  be a sequence with limit  $\alpha^*$  whose elements are in  $S_\delta(\alpha^*)$ . Define

$$A_\delta(\alpha^*) = \{x : \|x\| < 2(\sqrt{\alpha^*} + \delta)\}$$

—an open ball in  $\mathcal{X}$  with radius greater than twice  $\sqrt{\alpha^*}$ . Let  $\Xi$  denote the surface of the unit hypersphere in  $\mathbb{R}^p$ . For any  $x \in A_\delta^c(\alpha^*)$ —that is, any point such that  $\|x\|$  is greater than twice  $\sqrt{\alpha^*} + \delta$ —and  $\xi$  on the unit hypersphere

$$\|x - \xi\sqrt{\alpha_n}\| \geq \|x\|(1 - \sqrt{\alpha_n}/\|x\|) \geq \|x\|/2$$

since  $x$  cannot be closer to any point with radius  $\sqrt{\alpha_n}$  than it is to  $x\sqrt{\alpha_n}/\|x\|$ , the point with radius  $\sqrt{\alpha_n}$  on the common ray and since  $\sqrt{\alpha_n}$  is less than  $\sqrt{\alpha^*} + \delta$  by construction. Recall that

$$\tilde{p}(x|\alpha) = \int_{\Xi} (2\pi)^{-p/2} e^{-\frac{1}{2}\|x - \xi\sqrt{\alpha}\|^2} \pi(d\xi)$$

where  $\pi$  is the uniform distribution on  $\Xi$ . Let

$$g_1(x) = (2\pi)^{-p/2} \left[ I_{A_\delta(\alpha^*)}(x) + I_{A_\delta^c(\alpha^*)}(x) e^{-\frac{1}{8}\|x\|^2} \right]$$

Since  $\int g_1(x) \pi(d\xi) < \infty$  and  $g_1(x)$  dominates the integrand of  $\tilde{p}(x|\alpha_n)$ , we can say that

$$\lim_{n \rightarrow \infty} \tilde{p}(x|\alpha_n) = \tilde{p}(x|\alpha^*).$$

Furthermore,  $\int g_1(x) dx < \infty$  and  $g_1(x)$  dominates  $\tilde{p}(x|\alpha_n)$  by monotonicity of the integral. Therefore,

$$\lim_{n \rightarrow \infty} \int \tilde{p}(x|\alpha_n) dx = \int \lim_{n \rightarrow \infty} \tilde{p}(x|\alpha_n) dx = \int \tilde{p}(x|\alpha^*) dx.$$

Denote the Borel  $\sigma$ -algebra on  $[0, \infty)$  by  $\mathcal{B}$ , and choose  $B \in \mathcal{B}$ . Since the densities  $\tilde{p}$  and  $\tilde{q}$  are necessarily nonnegative and measurable, Fubini says

$$\tilde{R}(B|\alpha) = \int_B \int_{\mathcal{X}} \tilde{Q}(d\beta|x) \tilde{P}(dx|\alpha) = \int_{\mathcal{X}} \tilde{p}(x|\alpha) \tilde{Q}(B|x) dx.$$

Similarly,

$$\int_B \beta \tilde{R}(d\beta|\alpha) = \int_B \beta \int_{\mathcal{X}} \tilde{Q}(d\beta|x) \tilde{P}(dx|\alpha) = \int_{\mathcal{X}} \tilde{p}(x|\alpha) \int_B \beta \tilde{q}(\beta|x) d\beta dx.$$

Let

$$f(x) = \tilde{p}(x|\alpha^*)\tilde{Q}(B|x) \quad \text{and} \quad f_n(x) = \tilde{p}(x|\alpha_n)\tilde{Q}(B|x)$$

so that

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Since  $\tilde{Q}(B|x)$  is a probability,

$$f_n(x) \leq \tilde{p}(x|\alpha_n) \leq g_1(x).$$

Hence, by the dominated convergence theorem,

$$\lim_{n \rightarrow \infty} \int f_n(x) dx = \int \lim_{n \rightarrow \infty} f_n(x) dx = \int f(x) dx.$$

That is,

$$\lim_{n \rightarrow \infty} \tilde{R}(B|\alpha_n) = \tilde{R}(B|\alpha^*).$$

Therefore,  $\tilde{R}(B|\alpha)$  is a continuous function of  $\alpha$ .

Now note that

$$\int_B \beta \tilde{q}(\beta|x) d\beta \leq \int \beta \tilde{q}(\beta|x) d\beta.$$

By Proposition A.7 of Eaton *et alia* [11], for some bounded  $\psi(\|x\|^2)$ ,

$$\int \beta \tilde{q}(\beta|x) d\beta = -p + \|x\|^2 + \psi(\|x\|^2)$$

so that there exists a constant  $k > 0$  such that

$$\int \beta \tilde{q}(\beta|x) d\beta \leq k + \|x\|^2.$$

Let  $g_2(x) = k + \|x\|^2$ , and let  $g(x) = g_1(x)g_2(x)$ . Let

$$f_n(x) = \tilde{p}(x|\alpha_n) \int_B \beta \tilde{q}(\beta|x) d\beta \quad \text{and} \quad f(x) = \tilde{p}(x|\alpha^*) \int_B \beta \tilde{q}(\beta|x) d\beta$$

Note that

$$\lim_{n \rightarrow \infty} f_n(x) = f(x)$$

and  $g(x)$  dominates  $f_n(x)$ . One can show that

$$\int g(x) dx < \infty$$

by expanding the product  $g_1(x)g_2(x)$  and integrating the components. Now, by the Dominated Convergence Theorem,

$$\lim_{n \rightarrow \infty} \int f_n(x) dx = \int \lim_{n \rightarrow \infty} f_n(x) dx = \int f(x) dx.$$



That is,

$$\lim_{n \rightarrow \infty} \int_B \beta R(d\beta|\alpha_n) = \int_B \beta R(d\beta|\alpha^*).$$

Therefore,  $\int_B \beta R(d\beta|\alpha)$  is continuous as a function of  $\alpha$ . Since  $B$  is any Borel set, the continuity holds for both  $[0, m]$  and  $[0, \infty)$ . Finally, since  $T([0, m]|\alpha)$  is an algebraic combination of continuous functions, it is itself a continuous function.

**D.3. Moment conditions.** We are interested in moments of the transition kernel about the current state:

$$m_k(\alpha) = \int (\beta - \alpha)^k \tilde{T}(d\beta|\alpha).$$

We can express these moments in terms of  $\tilde{R}$  as

$$m_k(\alpha) = \frac{\int (\beta - \alpha)^k (\beta + \alpha) \tilde{R}(d\beta|\alpha) + c \int (\beta - \alpha)^k \tilde{R}(d\beta|\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}.$$

We know from Appendix A of Eaton *et alia* [11] that

$$(38) \quad \int \beta \tilde{R}(d\beta|\alpha) = \alpha + \phi_1(\alpha),$$

$$(39) \quad \int \beta^2 \tilde{R}(d\beta|\alpha) = \alpha^2 + 8\alpha + \phi_2(\alpha), \text{ and}$$

$$(40) \quad \int \beta^3 \tilde{R}(d\beta|\alpha) = \alpha^3 + 24\alpha^2 + \phi_3(\alpha).$$

where, as  $\alpha \rightarrow \infty$ ,  $\phi_1(\alpha) = O(\alpha^{-1})$ ,  $\phi_2(\alpha) = O(1)$ , and  $\phi_3(\alpha) = O(\alpha)$ . A similar argument shows that

$$(41) \quad \int \beta^4 \tilde{R}(d\beta|\alpha) = \alpha^4 + 48\alpha^3 + \phi_4(\alpha)$$

where  $\phi_4(\alpha) = O(\alpha^2)$  as  $\alpha \rightarrow \infty$ .

Let  $g_0(z) = (a + z)^{-p/2}$ , and let  $t_k(y) = E[g_0(U)U^k|y]$  with  $U \sim \chi_p^2(y)$ . Proposition A.2 of Eaton *et alia* [11] establishes that if  $Y \sim \chi_p^2(\alpha)$ , then

$$\int \beta^k \tilde{R}(d\beta|\alpha) = E \left[ \frac{t_k(Y)}{t_0(Y)} \middle| \alpha \right] < \infty.$$

Let

$$w_k(n) = \int_0^\infty \frac{g_0(z/2)^{n+p/2+k-1} e^{-z/2}}{2\Gamma(n+p/2)} dz$$

and note that  $E[g_0(U)U^k|y] = 2^k E[w_k(N)|y]$  where  $N|y \sim \text{Poisson}(y/2)$ . This last equality follows from expressing  $U$  as a Poisson mixture of  $\chi_p^2$  random variables. From the definition of  $t_k$  and  $w_k$ , we have that  $r_k(y) = 2^k E[w_k(N)|y]$ . Therefore,

$$\int \beta^k \tilde{R}(d\beta|\alpha) = 2^k E \left[ \frac{E(w_k(N)|Y)}{E(w_0(N)|Y)} \middle| \alpha \right].$$

Proposition A.7 of Eaton *et alia* [11] establishes that

$$\frac{E[w_k(N)|y]}{E[w_{k-1}(N)|y]} = \frac{y}{2} + 2(k-1) - \frac{p}{2} + \psi_k(y)$$

where  $|\psi_k(y)| \leq d_k/y$  for some finite positive constant  $d_k$ . This allows us to evaluate the right hand side of as

$$\int \beta^k \tilde{R}(d\beta|\alpha) = 2^k E \left[ \frac{E(w_k(N)|Y)}{E(w_{k-1}(N)|Y)} \cdots \frac{E(w_1(N)|Y)}{E(w_0(N)|Y)} \middle| \alpha \right].$$

Furthermore,  $E[\psi_k(Y)|\alpha] = O(\alpha^{-1})$  by Proposition A.8 of Eaton *et alia* [11]. Note that the third moment of a non-central  $\chi_p^2(\alpha)$  is  $\alpha^3 + O(\alpha^2)$  and the fourth moment is

$$\alpha^4 + 24\alpha^3 + 4p\alpha^3 + O(\alpha^2).$$

Combining these results leads to (41).

We now begin to find expressions for the transitional moments  $m_k$  in terms of equations (38), (39), (40), and (41) as  $\alpha$  becomes large. First, note that if  $0 < c < \infty$ , then

$$c \int (\beta - \alpha)^2 \tilde{R}(d\beta|\alpha) = O(\alpha) \quad \text{and} \quad c \int (\beta - \alpha)^3 \tilde{R}(d\beta|\alpha) = O(\alpha^2).$$

Express the first transitional moment as

$$\begin{aligned} m_1(\alpha) &= \frac{\int (\beta - \alpha)(\beta + \alpha) \tilde{R}(d\beta|\alpha) + c \int (\beta - \alpha) \tilde{R}(d\beta|\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)} \\ &= \frac{\int \beta^2 \tilde{R}(d\beta|\alpha) - \alpha^2 + c \int \beta \tilde{R}(d\beta|\alpha) - c\alpha}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)} \\ &= \frac{8\alpha + O(1)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}. \end{aligned}$$

Since  $(\beta - \alpha)^2(\beta + \alpha) = \beta^3 - \beta^2\alpha - \beta\alpha^2 + \alpha^3$ , we have that

$$\begin{aligned} \int (\beta - \alpha)^2(\beta + \alpha) \tilde{R}(d\beta|\alpha) &= \alpha^3 + 24\alpha^2 + O(\alpha) - \alpha^3 - 8\alpha^2 - O(\alpha) - \alpha^3 - O(\alpha) + \alpha^3 \\ &= 16\alpha^2 + O(\alpha). \end{aligned}$$

Express the second transitional moment as

$$\begin{aligned} m_2(\alpha) &= \frac{\int (\beta - \alpha)^2 (\beta + \alpha) \tilde{R}(d\beta|\alpha) + c \int (\beta - \alpha)^2 \tilde{R}(d\beta|\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)} \\ &= \frac{16\alpha^2 + O(\alpha)}{\int (\beta + \alpha + c) \tilde{R}(d\beta|\alpha)}. \end{aligned}$$

Since  $(\beta - \alpha)^3(\beta + \alpha) = \beta^4 - 2\beta^3\alpha + 2\beta\alpha^3 - \alpha^4$ , we have that

$$\begin{aligned} \int (\beta - \alpha)^3(\beta + \alpha) \tilde{R}(d\beta|\alpha) &= \alpha^4 + 48\alpha^3 + O(\alpha^2) - 2\alpha^4 - 48\alpha^3 - O(\alpha^2) \\ &\quad + 2\alpha^4 + O(\alpha) - \alpha^4 \\ &= O(\alpha^2). \end{aligned}$$

Since our expressions for the transitional moments all share a common denominator, the ratio  $m_3(\alpha)/m_2(\alpha)$  may be evaluated as

$$\begin{aligned} \frac{m_3(\alpha)}{m_2(\alpha)} &= \frac{\int (\beta - \alpha)^3(\beta + \alpha) \tilde{R}(d\beta|\alpha) + c \int (\beta - \alpha)^3 \tilde{R}(d\beta|\alpha)}{\int (\beta - \alpha)^2(\beta + \alpha) \tilde{R}(d\beta|\alpha) + c \int (\beta - \alpha)^2 \tilde{R}(d\beta|\alpha)} \\ &= \frac{O(\alpha^2)}{16\alpha^2 + O(\alpha)} \\ &= O(1). \end{aligned}$$

## REFERENCES

- [1] CHUNG, K. L. and FUCHS, W. H. (1951). On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6** 1–12.
- [2] BERGER, J. and STRAWDERMAN, W. E. (1996). Choice of hierarchical priors: admissibility of normal means. *Ann. Statist.* **24** 931–951.
- [3] BERGER, J., STRAWDERMAN, W. E. and TAN, D. (2005). Posterior propriety and admissibility of hyperpriors in normal hierarchical models. *Ann. Statist.* **33** 606–646.
- [4] BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–903.
- [5] EATON, M. L. (1982). A method for evaluating improper prior distributions. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) Academic Press, Inc., New York.
- [6] EATON, M. L. (1992). A statistical diptych: Admissible inferences—recurrence of symmetric Markov chains. *Ann. Statist.* **20** 1147–1179.

- [7] EATON, M. L. (1997). Admissibility in quadratically regular problems and recurrence of symmetric Markov chains: Why the connection? *J. Statist. Plan. and Infer.* **64** 231–247.
- [8] EATON, M. L. (2001). Markov chain conditions for admissibility in estimation problems with quadratic loss. In *State of the Art in Probability and Statistics: Festschrift for Willem R. van Zwet* (M. de Gunst, C. Klaasen, and A. van der Vaart, eds.) 223–243. *IMS Lecture Notes Ser.* **36** IMS, Beechwood, OH.
- [9] EATON, M. L. (2004). Evaluating improper priors and recurrence of symmetric Markov chains: an overview. In *A Festschrift for Herman Rubin* (A. DasGupta, ed.) 5–20. *IMS Lecture Notes Ser.* **45** IMS, Beechwood, OH.
- [10] EATON, M. L., HOBERT, J. P. and JONES, G. L. (2007). On perturbations of strongly admissible prior distributions. *Ann. Inst. H. Poincaré Probab. Statist.* **43** 633–653
- [11] EATON, M. L., HOBERT, J. P., JONES, G. L., and LAI, W.-L. (2008). Evaluation of formal posterior distributions via Markov chain arguments. *Ann. Statist.* **36** 2423–2452
- [12] HOBERT, J. P. and ROBERT, C. P. (1999). Eaton’s Markov chain, its conjugate partner, and  $\mathcal{P}$ -admissibility. *Ann. Statist.* **27** 361–373.
- [13] HOBERT, J. P. and SCHWEINSBERG, J. (2002). Conditions for recurrence and transience of a Markov chain on  $\mathbb{Z}^+$  and estimation of a geometric success probability. *Ann. Statist.* **30** 1214–1223.
- [14] HOBERT, J. P., TAN, A. and LIU, R. (2007). When is Eaton’s Markov chain irreducible? *Bernoulli* **13** 641–652.
- [15] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* 1361–380. University of California Press, Berkeley.
- [16] JOHNSON, B. W. (1991). *On the Admissibility of Improper Bayes Inferences in Fair Bayes Decision Problems*. PhD thesis, University of Minnesota.
- [17] KASS, R. E. and WASSERMAN, L. (1996). The selection of prior distributions by formal rules. *J. Amer. Statist. Assoc.* **91** 1343–1370.
- [18] LAI, W.-L. (1996). Admissibility and recurrence of Markov chains with applications. Ph.D. thesis, University of Minnesota.
- [19] MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [20] REVUZ, D. (1984). *Markov Chains*, 2nd ed. North-Holland, Amsterdam.

- [21] SCHERVISH, M. J. (1995). *Theory of Statistics*. Springer.
- [22] STEIN, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.
- [23] TARALDSEN, G. AND LINDQVIST, B. H. (2010). Improper priors are not improper. *The Amer. Statist.* **64** 154–158.